

消費者委員会 人工知能（AI）技術の利用
と消費者問題に関する専門調査会（第7回）
議事録

内閣府消費者委員会事務局

消費者委員会 人工知能（AI）技術の利用と消費者問題に関する
専門調査会（第7回）
議事次第

1. 日時 令和8年7月30日（木）10:00～11:01
2. 場所 消費者委員会会議室及びテレビ会議
3. 出席者

（委員）

【会議室】

小塚座長、丸山座長代理、大塚委員、岡崎委員、坂下委員、野村委員、
馬籠委員

【テレビ会議】

加藤委員、唐沢委員、

（オブザーバー）

【テレビ会議】

鹿野委員長、大澤委員、善如委員、山本委員

（事務局）

小林事務局長、吉田審議官、友行参事官、江口企画官

4. 議事

（1）開会

（2）消費者問題としての自律的意思決定の阻害に関する整理

唐沢委員プレゼンテーション

（3）閉会

《1. 開会》

○小塚座長 皆様、おはようございます。定刻になりましたので始めさせていただきます。

こちらは大変暑い東京でございますけれども、お忙しい中、お集まりいただきましてありがとうございます。

「消費者委員会人工知能（AI）技術の利用と消費者問題に関する専門調査会」第7回会合になります。ただいまから開催いたします。

本日の出席状況でございますが、座長代理の丸山座長代理と大塚委員、岡崎委員、坂下委員、野村委員、馬籠委員、そして、私、小塚が会議室におります。それから、テレビ会議システムのほうで加藤委員、唐沢委員に御出席をいただいております。

田中委員は所用により御欠席という御連絡をいただいております。

それから、オブザーバーの消費者委員会委員の先生方ですが、本日は鹿野委員長、大澤委員、柿沼委員、善如委員、山本委員にテレビ会議システムにて御出席いただいているということです。

それでは、いつものことですが、会議の進め方につきまして事務局から御説明ください。よろしくお願いします。

○江口企画官 本日はテレビ会議システムを活用して進行いたします。一般傍聴者にはオンラインにて視聴いただき、報道関係者のみ会議室にて傍聴いただいております。

議事録につきましては、後日、消費者委員会のホームページに掲載いたします。議事録が掲載されるまで、ユーチューブでの見逃し動画配信を行います。

配付資料につきましては、お手元の議事次第に記載してございます。もし不足の資料がありましたら、事務局までお申し出くださいますようお願いいたします。

以上です。

《2. 消費者問題としての自律的意思決定の阻害に関する整理

唐沢委員プレゼンテーション 》

○小塚座長 ありがとうございました。

それでは、本日の議事は「消費者問題としての自律的意思決定の阻害に関する整理」ということで、唐沢委員からプレゼンテーションをいただくということになっております。そして、その後で質疑応答、意見交換をしてまいります。

唐沢委員、御用意はできましたでしょうか。20分程度でお話しいただけると承知しておりますが、よろしいですか。

○唐沢委員 よろしく申し上げます。本日出張中なので、資料共有も事務局のほうにお願い

いしておりますので、よろしくお願いします。

それでは、お話をさせていただきます。

唐沢と申します。本日はよろしくお願いいたします。

報酬環境と社会的な側面ということから、消費者とAIの問題について本日お話をさせていただきます。

次のスライドをお願いいたします。

本日は、事務局からの御依頼事項として、こちらの3つです。報酬、強化学習に関しての心理学の話、それから、これまで挙がった問題の心理的な過程についてということ、あと、認知コントロールについて可能であればという御依頼をいただきました。

これらを踏まえまして、本日は大まかに2つの話をさせていただきます。

1つ目は報酬と強化学習に関する話で、これは心理学の話です。強化学習の話をさせていただきますが、ポイントは、これまでの議論の中では、人の心理について意思決定という点から消費行動を考えてきた、つまり、意思決定のプロセス、好みを形成したり、私はこれらの中からこちらを選ぶという主体的な意思決定者としての人間というものに焦点を当てた議論というのがなされてきたと思うのですが、強化学習というのはそれとは違う人間観を提供するものです。具体的にはどういうものかというのはこれから説明しますが、私たちの行動というのは環境がもたらす報酬や罰の与えられ方によって決められてしまうというニュアンスを含みますので、このことを前提として聞いていただければと思います。

2つ目は、私自身、社会心理学者ですので、強化学習ということを前提に社会的な側面にも着目して、消費者やAIをめぐる心の働きということについて話をさせていただければと思いますので、よろしくお願いいたします。

では、次のスライドをお願いします。

強化学習が心理学の中でどういうものかということについての位置づけのスライドになります。心理学の中で割と古い研究テーマなのですけれども、こちらに例えばソーンダイクとかスキナーとかという人物名が挙がっていて、どんなことを行ったかを記載しています。時間も限られていますので簡単に話をさせていただくと、もともと動物を使って、動物が経験に基づいてどのように行動を変えていくのかを観察し、記述し、また、検討してきた研究の流れになります。動物を箱に入れて何か行動をすると例えば餌がもらえるなどの場面で、行動を観察していくわけですが、餌の与えられ方、もしくは罰を与えることもありますが、それらによって行動の形づくられ方、動物がもともと持っていなかった行動のレパトリーを獲得していくことを検討する研究になります。

ですので、強化学習の研究の流れや考え方は、下のほうに書いてあるのですけれども、行動が私たちの意図とか意思で決まるというよりは、その行動を行ったらどんな結果がもたらされたか、それによってつくられていくのだというような行動観、まだ、動物を対象とした研究が多いのですが、人間もそういう側面がありますから、人間観や理解の基礎になっている領域になります。

別の言い方をしますと、報酬とか強化という言葉を使って解説していくわけですが、私たちがやる行動を獲得する。それを繰り返すということが、何か納得してやるとか、こういう意思を持っているからというよりも、これを行ったからどんな結果が得られたのか、自分の経験の履歴によって決まるという、行動と結果の関係から説明するパラダイムだという前提があります。

では、次のスライドをお願いします。

強化学習のメカニズムというのは単純なものです。真ん中の基本過程のところを見ていただくとよいのですが、何か行動を行います。当初は探索的に行うことが多いですが、だんだん学習が進むにつれて、餌をもらえる行動を繰り返すなどがみられるわけですが、ある行動をやったときにそれは何かの結果をもたらします。餌が来るとか、いいことがあるとか、もしくは罰が来るとかということがある。動物であれ、人間であれ、行う主体の持っている欲求、また、行動がもたらす結果に対する期待があり、それらに基づいて、結果の予測があるわけですが、それとの差を評価して、プラスの価値を持つのか、マイナスの価値を持つのかということが決まります。プラスの価値を持つものというのは、その行動について、今後、行動そのものにプラスの得点、値が付くイメージで、行動がもたらし得る価値がプラスになり、罰などマイナスの結果がもたらされるものはマイナスの行動価値が付与されるとなります。

ある行動について数値がついているというイメージを持っていたらいいかと思うのですが、複雑なことではなく、よい結果が来るような行動、しかも、それが自分の持っている期待とか欲求という観点からよりよいものが来るものは行われやすくなって、より嫌な結果がもたらされるものは行われにくくなるという、単純なメカニズムで出来上がっているのが強化学習の基本原理になります。

また強化学習については、左側にある3つの用語、強化子、予測誤差、強化スケジュールが使われます。予測誤差は今説明したとおりなので、ここでとどめますが、強化子と強化スケジュールでは後で説明をします。

右側の強化学習が結局何をもたらすかということも、実際の日常経験から既によく御存じのことだと思うのですが、強化のプロセスによってある特定の行動が取られやすくなり、繰り返しやすくなり、それが継続するとか習慣化していく。自動化していくということも起これば、特定の行動が固定されるということも起こります。

この行動の固定化のところで一言申し上げておくと、すぐに報酬が来る行動としばらく待たないと報酬が来ない行動を比べると、すぐに来るもののほうがより固定化する、選ばれやすくなる傾向があります。ダイエットの例ですと、甘いものを食べる行動はすぐ報酬が得られますが、それは健康によくない。食べずにいると健康が守られるというのは、より後に報酬が来ることになるので、甘いものを食べる行動と我慢するという行動を比べると、どうしても食べる行動が起こりやすい。すぐ報酬が来る行動に、傾きやすいのが特徴の一つです。

まとめますと、下のほうに「強化学習とは、行動の結果と予測の差をもとに、その行動を再び選ぶ価値を更新していく過程」と書いてあるのですが、平たく言うと、よい結果をもたらす行動をより行いやすくなる。ネガティブな結果をもたらす行動より行いにくくなるというような基本的なプロセスで構築されているものを強化学習と呼ぶと御理解ください。

では、次をお願いします。

ここからは言葉の説明になって恐縮なのですが、報酬とか罰とか強化子という言葉の話を最初にします。報酬とか罰は一般的な言葉で、これはそのとおり御理解いただけたと思うのですが、強化子という言葉は心理学の用語ですので、ここで確認させてください

報酬や罰というのは一般的によいものとか悪いものと理解できますが、強化子というのはその中でも特に効果を持つもの、効果を持つ報酬とか罰のことを強化子と呼びます。日常の理解では報酬や罰でも問題ないのですが、この言葉を使ってスライドの説明を記載していますので、ここで確認をしているという次第です。

強化子、報酬や罰ですね。その次の下のところなのですが、これらには大まかに2つあって、一つは一次性、もう一つは二次性と呼ばれるもので、一次性は生得的な価値を持つような食べ物とか水とか絶対必要なもの、生き物として必要なもの、価値を持つものが一次性。二次性というのは、本来、それ自体は生得的な価値を持たないのですが、経験とか社会的に構築されていくことによって価値を持つようになっていくもの、お金などを二次性と呼びます。

なので、ここでのポイントは、単に身体的な快、不快以外にも、特に人間にとっては二次性の強化子が非常に重要な強化子となって、行動をシェイプしていく効果があることです。

なので、強化子とは人の行動を方向づけたり、繰り返させる、頻度を上げたり下げたりするような効果を持つものです。ただ、何がその人の行動を変える強化子になるかは個人の状態とか文脈とか経験によって変わる。個人差があることもポイントになると思います。

その次のスライドをお願いします。

正の強化とか負の強化、弱化という言葉についてもお話しします。

弱化という言葉が一般的ではないと思うのでリストしておきましたが、要するに罰が弱化に相当します。行動を変容させるための強化子の与え方に4つのパターンがあり、1つ目が正の強化と呼ばれるものです。行動の結果良いことが得られるのがこちらに相当します。

2つ目が、行動の結果嫌なことがなくなった場合、それもその行動を増加させますが、これが負の強化と言われます。

3つ目が正の弱化ということで、いわゆる典型的な罰として皆さん理解していただいていることなのですが、行動した後に嫌なことが来るとその行動が減るとというのが正の弱化。

それから、行動した後、よかったものがなくなってしまうというのが負の弱化になります。

す。

これらはそれぞれ例を挙げておきましたので、個別に見ていただければと思います。強化子の与え方のパターンはこの4つになりますが、さらに、どのようなスケジュールで与えるかが行動の頻度とか行動のパターンを決めていく点が、強化学習の一つの重要なポイントになりますので、それについて次のスライドで説明します。

次をお願いいたします。

今、強化の与え方、罰とか報酬の与え方はいろいろあることを話したのですが、どんなスケジュールに与えるかということも強化学習では大事です。大まかにこの4つがあります。行動の回数で決まるもの、時間間隔で決まるもの、それから、固定化したものと変動していくものというパターンで4つのバリエーションができています。

一番左上の固定比率というのは、一定の反応ごとに強化していくものです。例えば5回買うと1回タダになるなどが相当しますが、これを行ったときの行動の特徴というのは、反応率は一定維持されるのですけれども、1回無料になったそのすぐ後、強化の直後は、反応が鈍くなるのですけれども、回数がかさなり、5回に近づくにしがって加速するカーブになります。5回買うと1回ただのときに、4回買ったなら、次の5回目は急いで買ってしまうみたいなイメージです。

その下の固定間隔というのは、一定の経過時間ごとの最初の反応で強化されるもので、実際の事例は難しいのですが、タイムセールがこれに相当するのかなと思います。3か月に1回セールがあると分かっていると、その時間になると反応が増加するようなことが固定間隔のスケジュールになります。

右上の変動比率は非常に高い反応を示すことで知られている強化のバリエーションで、平均すると一定の回数ごとなののですけれども、必要な反応数自体は毎回変動するもので、ギャンブルが典型的な例です。

その下の変動間隔は、時間間隔で決まるもので、平均すると一定時間なのですが、どのくらいの時間で強化させるかが変わってくる。これも事例が難しいのですが、不定期に届くクーポンみたいなものなのかなと思います。中程度に安定した反応が続くもので、例えばこまめにチェックするとか習慣化していくような反応を引き起こします。

これらのスケジュールにより、行動が操作されることになります。

次をお願いします。

ここまでの学習心理学の基本的な話をさせていただいたのですが、ここからはAIとの関連を考えたいと思います。

今までお話ししたような強化の仕方とか強化スケジュールというのは、これまでも消費者に対して用いられてきたわけですが、AIにより何が変わるかがポイントになると思います。

2つ挙げたのですけれども、一つは、AI自体が個人の反応に基づいて、強化のスケジュールやタイミング、強化の仕方について、効果的に設計することで、AIが行動変容を非常に

促しやすい環境をつくることです。個別最適化された報酬（強化子）をうまくつくり、行動変容を量的に加速していくことが可能になりました。

2つ目は、AIとの相互作用を消費者が行う中で、消費行動に影響する働きかけを行うことです。例えばAIが相互作用を通して、消費者の状況、状態を推定して、それに基づいてより効果的なメッセージを送ることもできます。また、個別最適化された、相談とか対話、もしくは明示的な相談、対話ではなくても、AIとの相互作用そのものも社会的な要求を満たすものになるので、非常に有効な消費の対象となっていくことになります。このような社会的要求の話は後に補足します。

それから、相談の中で選択の意味と理由そのものをAIが構築してくれる。そのことで質的な変化をもたらすといえますか、AIが選好を形成する主体として位置づけられるようなポジションを持つことになります。ただ、これは言葉遣いが難しいと思うのは、AIが私たちの選択を助けるとも表現できますし、良い助けもあるので、ネガティブなこととして捉えるべきというものではない。何がよいか、何が悪いかを見極めないといけないのですけれども、この見極めは実際、結構難しい問題かもしれないと思っております。

次をお願いします。

今、相談ということを申し上げたのですけれども、相談そのもの、AIと対話して相談すること自体が、消費行動でもあることを示す、その点を解説するのがこのスライドになります。

負の強化、回避学習という言葉が出てきておりますが、内容は単純で、ある行動hにより嫌なことが取り除かれると、その行動が起こりやすくなるというのが負の強化ですが、それに基づき、回避学習が行われます。つまり、嫌なことが起こる前にそれを予測して、避けるための行動を私たちは学習しますよね。例えば、研究としては、その下にあるような、音が出ると少し遅れて電気ショックが来る場面で、音が出たら逃げるという行動を動物が学習するなど、単純なパラダイムが例になります。私たちは大きな不快ではなくても、少しの不安や後悔する可能性を予測すると、それを避けるための行動を行おうとしますが、その固定化、習慣化が、AIとの相談によりつくられるのがこのスライドのポイントになります。

私たちは物を選ぶときに、あまり考えずに買うこともあります。そのあと、もっと良いものがあつたのではないかと思ったりもします。ですから、調べてから買おう、調べずに買うのは良くないと思う場面は結構多いと思いますし、そういうことになると、物を買うという場面自体が、不安という強いかもしれませんが、調べないと後で後悔するかもというネガティブなことを予期させる場面にもなってきます。そういう場面では人に相談するのでもいいのですけれども、AIに相談して、それでうまく決めることができると、不安や後悔を回避することができるということで強化されます。これが定着していくと、購入場面そのものが、相談しないとまずい場面、脅威場面化してしまい、不安や後悔を回避するために、AIに相談することが固定化してしまうことが起こり得ます。

こうなると、消費行動自体は、物を買う行動なのですけれども、それ以前のAIと相談する、口コミなどを確認することも含みますし、むしろそれ自体が消費行動のメインになり、安心するための相談自体が消費される対象となっていくということが起こり得ます。

次のスライドをお願いします。

相談するのは悪いことではないのですけれども、AIの相談というのはいろいろ問題となることを含んでいます。一つは、人に相談するのとは異なり、歯止めがかからない。何度聞いても大丈夫だし、相手の負担を考える必要がない、社会的なコストがかからないということがあります。歯止めがなく、いつでもどこでも応答してくれるAIの特性により、物を買うときの後悔を避けるためにはAIに相談することが必要だという信念から逃れることを難しくします。相談しなくて大丈夫だと分かるためには、相談しなくてうまくいった事例を経験しないといけないのですけれども、それ自体が難しくなるのです。

AIへの相談が繰り返される安全行動、これをすれは大丈夫だと思える行動として機能し、不安がAIに相談したら解消されると、AIを使うこと自体が維持・強化される、このようなメカニズムをAIへの相談が持つてしまうことがあります。

AIに相談することには、良いことももちろんあるのですけれども、こういうリスクがあることもポイントになるかなと思います。

その次のスライドをお願いします。

もっとも、消費行動自体が脅威や不安、後悔をそれほどもたらすものではないという反論もあると思います。それはそのとおりだとしても、やはり、AIを使う行動自体、行動を生じさせる条件として言われていることが非常に整っている環境であるのも事実です。ある行動が行われやすい環境、状況というか、条件は、大まかに3つあることが言われており、十分な動機づけと、その行動がどれだけやりやすい、簡単か、やる能力があるかということ、それから、きっかけの3つがあります。スライドに書いてあるように、AIを使う行動は、問題解決や好奇心を満たすという動機づけが十分私たちの中で喚起されやすいですし、負担も少ない、いつでもできるという簡単さも持っていますし、きっかけとしてもいろいろな組み込みとか通知で、使うことを誘発する仕組みを入れることができるので、これらがそろっています。AIを使うという行動は非常に習慣化しやすい行動の一つなのです。

また、特に生成AIが入ってくると、相互作用を通じてそれが動的に調整されていく可能性が高いので、消費者の状態を推定してくれて、意味のある答えをつくってくれる。そうすると、このAIは私をちゃんと分かってくれるよい存在、よいパートナーみたいに感じることもなるので、社会的な動機と呼ばれるものを満たすことで、より利用が促進される可能性もありますし、また、利用履歴をもとにAIが先回りしていろいろ行うことを踏まえますと、利用者のコストを減らして行動を容易化していくので、消費行動としてのAI利用が大変起こりやすい行動だということになります。

次のスライドをお願いします。

消費行動として起こりやすいとか、AIとの相互作用について話をしましたが、AIとの相

相互作用については既にいろいろ議論がされている状況で、AIが単に道具ではなくて、我々にとって何者かということを考えたときに、例えば友達としてAIを捉えるというAI観を提供している研究者の方もおられますし、パートナー、コンパニオンとしてAIを捉えるという論調の議論もあります。

そのことを踏まえて考えますと、消費が持つ社会的な意味を確認しておく必要があるかなと思います。物を買って必要を満たすということが消費行動の主な意味だと思うのですが、それ以外にもいろいろな社会的な意味を持っているということは古くから言われていて、こちらにあるような自己表現とか、自分が何者であるか、所属を明らかにするとか、他者からの承認を消費によって得るとか、多くの人が選んでいるものを自分が使うことで、私は正しい、大丈夫なのだとことを確認するとか、このような社会的な意味があることが言われています。これら社会的な意味は、私たちの社会的な動機とつながっているわけです。例えば、自己表現をしたい、自分が何者かを確認したいという動機とつながっているのですけれども、これら社会的な動機を満たす消費が、AIとの相互作用によってより満たされることが起こりえます。

社会的な意味が、他者からの働きかけでより強固になること、こちらの右側にあるいろいろなメッセージで強化されていくことは、私たちは経験していることで、店舗でこういうことを言われますから、AIだけの話ではないのですけれども、これら社会的な意味を満たす消費を促進する働きかけというのは、AIがパーソナルに働きかける親身なコミュニケーションを行う、必要とするときにはすぐに反応する、個人に最適化した反応をするなどにより、AIと人の間の信頼関係を作ってしまう、信頼できるエージェントとしてAIが、より大きな影響力を持ってしまう結果にもつながります。

これは消費行動だけではないので、そもそも私たちにとってAIが何者かというコンテキストの中で、人とAIとの関係を考えなければならない、大きな文脈の中でも議論が必要かなと思います。道具としてのAIであるなら、こういうときはよくない道具になるから使うのをやめましょうということが言えたとしても、AIがパートナー、友達として捉えられている状況の中で、AIには注意してくださいという働きかけは、あなたにとって有害な友達だから付き合うなということと、同様のメッセージを含むかもしれません。そうすると、AIに対して注意すべきというメッセージをどう伝えるかにも、留意が必要になると思います。

次をお願いします。

ここまでAIとの話をしましたが、最後に認知コントロールということで事務局からいただいたので1枚だけ作りました。認知コントロールは何かはここに書いてあるとおりののですが、端的に言うと立ちどまって考えるということになります。ちょっと時間を置いて考えましょうとか、いきなりやらないで、ということなのですが、コントロールするときには大まかに3つの要素があります。一つは衝動的・自動的な反応を止めること。2つ目が何を自分が目標としているとか、判断の基準をもう一回ちゃんと考え直すこと。3つ目がほかの見方とかオルタナティブの方に意識を向けることです。

このそれぞれに対して何ができるかというのが右側のほうに書いてあって、これは言わば認知コントロール、自衛策のハウツーでもあるので、こういうことに注意しましょうとは言えます。これらに注意することによって、AIがつくった行動選択環境といいますか報酬環境から一步離れて、自分の目標とか基準を見直して再評価しようというメッセージは送れると思います。ただ、AIの働きかけが継続的で個人化されて、実はその影響をどれだけ自分が被っているかは結構見えにくいなかで、個人の注意力とか抑制力でどこまで抵抗できるかという、これで大丈夫ですとは、やはり言いにくいと思います。

これまでいろいろな議論がなされており、ここでも設計とかAIの在り方をどうしたらいいのかは議論されてきたと思うのですが、報酬環境や強化学習という文脈を踏まえるならば、そのことを前提として、AIの在り方、どのように活用するかという外部の環境設計のほうにいろいろな提言、実際に何をすべきかを盛り込む必要があると思います。また、それをどのように効果的な運営のノウハウ、ルールにつなげていくのかということも非常に重要なことなのではないかなと思っております。

最後のスライドはこれまでの文献を示したもので、私の話としてはここまでになります。どうもありがとうございました。

○小塚座長 唐沢委員、御出張先からありがとうございました。

それでは、どなたからでも御質問、御意見をいただきたいと思います。会場の方は挙手で、それから、オンラインの方はチャットでお願いします。

では、野村委員、どうぞ。

○野村委員 唐沢委員、ありがとうございました。

今日のお話を伺って、AIに対する行為依存と関係依存は待ったなしで起こるということを経験的に御説明いただいたという解釈でよろしいでしょうか。

○唐沢委員 必ず起こるとは言えませんが、非常に起こりやすい環境が整っているということを経験学の知見を入れながらお話しさせていただいたのかなと思っております。

○野村委員 ありがとうございます。

あともう一点なのですが、最後のスライドを拝見して思ったのが、結局、単にメッセージを出すだけでは駄目という話があったのですが、つまるところ、やはり消費者側のAIに対するリテラシーの高低によって結果なり変わってしまうのかなということが想像されるのですが。

○唐沢委員 リテラシーは大変重要だと思いますし、リテラシーの個人差により、よりある人のほうが抵抗力が強いことは予測できます。だから、それを高めるという働きかけが大事だということはそのとおりですので、それを否定するつもりはないです。また、リテラシーにかかわらず、こういう問題がありますよねということを分かりやすく伝えていくことは大事だと思っておりますので、むしろ行うべき重要なことの一つです。ただ、それだけに頼るということが問題なのではないかなというのが言いたいことです。

○野村委員 ありがとうございました。

○小塚座長 ありがとうございます。

それでは、どなたからでもそのほか御意見。

丸山座長代理、お願いします。

○丸山座長代理 御報告ありがとうございました。大変勉強になりました。

私からは2点お伺いしたいのですが、1点目が8ページ以降に関わることでございまして、対応を要するネガティブな状況というものをどのように把握すべきかということを考えたときに、結局、選好が形成されて変化するというところに着目するよりも、むしろバイアスというか認識のゆがみや思考の偏りというものが形成されたりAIによって増幅されるといったことに着目したほうがよいのかという辺りをまずは第1点としてお伺いできればと思いました。

○唐沢委員 どちらも重要だと思います。何がネガティブかを決めるのは実は難しいという話をしたと思うのですが、要するに本人が納得しているという状況であっても、外部から見るとおかしいという状況は当然あり得ますし、バイアスというのも、例えば実験場面などでは、ある反応をバイアスだと言うのは可能なのですが、日常の判断であなたの判断はバイアスがかかっていますよねということを本当に確信を持って言えるかという、それも実はかなり難しいと思います。ただ、こういうところに問題が起こり得るかを明らかにする、どのようなバイアスとかがあり得るのかを丁寧にリストし、問題の可能性のマップとか見取り図みたいなものを示していくことによって、起こり得る問題に対応していくというやり方が考えられると思います。

個別に問題が起こったときには、臨床的といいますか、例えば問題があって相談に来られたときには、それに即して対応することになると思うのですが、それ以前に発見するとなると、可能性の中でどこに起こり得るかを予見することが必要なので、問題の可能性については、むしろ広めに扱うほうがよいのかなと思っております。

○丸山座長代理 ありがとうございます。大変よく分かりました。

2点目なのですが、生成AIと感のよい親身なセールスマンというのを比べたときに、その違いというのはやはり量的なものにとどまらず、御指摘いただいた歯止めがなくなるというのは質的な違いであると捉えてしまってよいのかというところを教えてくださいました。

○唐沢委員 量の違いは、あるところでは私は質的な違いに転換してしまうと思います。質なのか量なのかという議論は複雑な議論になってしまうのかもしれませんが、歯止めがないこと、例えばいつでも対応してくれるとか、コミュニケーションが途切れなくうまくなされること自体が、対象に対しての信頼とか、また依存という言葉が先ほど出てきましたが、そのような心的な状態につながる、また、心の状態として、ある閾値みたいなもの。自分の中で踏みとどまるラインを越えて、全面的に信頼してしまうのだということで、量の問題が質の変化につながる可能性がある。それがどこで起こるかは、個人差もあるでしょうし、見極めるのは難しいかと思うのですが、単に量的というよりも、

関係性の変容みたいなことですので、質的な変化といえるのではと思います。お答えになっているでしょうか。

○小塚座長 ありがとうございます。

もともと唐沢委員のAIには幾らでも頼ることができるというのは、AIに相談するという文脈でおっしゃったのですが、丸山座長代理の今の御質問はそれはむしろ勧誘事業者からの消費者に対しての働きかけの場面でも同じだという前提での御質問だったのですか。

○丸山座長代理 そういうものに使われていくときに問題が生じ得るだろうと考えたときに、量的なところだけではなくて質的な違いがあるのかを伺っておきたかったという趣旨でした。

○唐沢委員 すみません。理解不足で申し訳なかったです。

○小塚座長 いえ、唐沢委員の今日のお話を広げていただいたということで大変重要なポイントだと思います。ありがとうございます。

そのほかいかがでしょうか。

では、大塚委員、お願いします。

○大塚委員 ありがとうございます。

御報告ありがとうございました。特にAI利用における消費者の行動心理の理解が深まったように思います。

私からは、スライドでいうと10ページあたりのAIへの相談が不安の維持システムに変容するという点について質問いたします。唐沢委員の御発表では、こういった不安の維持システムに変容してしまうという点はネガティブに捉えられていたように思います。ただ、この不安というものが購買において失敗しないようにするためのものである、不安を抱くことによってよりちゃんと調べたりするというものだとなると、必ずしも不安を抱くこと自体はネガティブに捉える必要はないようにも思いました。例えば何らかの購買行動をするときに、必ずAIに相談する。AIがこういった選択肢もあるのではないのと提示して、それに対してそうかもねといって、別の購買行動を取る。こういうことを繰り返すことによって、むしろ全体としては後悔をしないような行動につながっていく。そういった可能性もあるように感じました。

そこで、唐沢委員がどこにネガティブを感じているのかというところを確認させていただきたいと思います。一つはこの不安の維持システムというか、購買に必ず不安を感じてしまうというそれ自体があまり望ましくないというお話なのか、あるいは、その結果、不安の維持システムとしてAIへの相談が出てくるといのが何かネガティブな結果をもたらすのか。例えば適切な購買行動が減ってしまうとか、あるいはAIの依存自体が必ずしも望ましくないであるとか、そういった結果を意識していらっしゃるのか、そして、その場合にはどういった結果が望ましくないと考えられているのかという点を伺えればと思います。

○唐沢委員 御質問ありがとうございます。

使うことでよりよい選択ができるというのはそのとおりですし、また、どんどん使って

いくことでうまく使い方というのは私たちは学習できるので、おっしゃるとおり、相談が駄目だよという話ではもちろんないわけです。

ただ、これがネガティブになり得るとすると、どのようなメカニズムなのかという話として御理解いただければと思うのですけれども、一つはAIを使うことでうまくいくという場面を繰り返していくことで、うまく使えるようになっていくのだけれども、使わなくても大丈夫というような状況を経験することを妨げるので、AIなしでやっていくということが、もしかしたら難しくなってしまう、何でもAIに聞くという傾向を強めてしまう可能性が出てしまう。この何でもAIという傾向は消費行動としての相談以外にも般化される可能性があります。AIを使うと消費行動でうまくいく経験は、ほかの場面でもAIに頼めば何でもしてくれるという認識、そういうことの学習につながるので、AI依存がより深まる問題になると思います。

よいのか悪いのかということについては、うまくAIを活用する、どんな場面でもうまく使える存在として自分を鍛えらることも表現できるので、必ずしも悪いということではないと思います。しかしながら、AIがないと不安になってしまうことが起こり得るかもしれないとすると、その意味ではネガティブになる。何がネガティブかについては、AI依存、またそれにつながる行動が強化される点が挙げられると思います。

○大塚委員 ありがとうございます。

○小塚座長 ありがとうございます。

そのほか。

では、坂下委員、それから、馬籠委員の順番でお願いします。

まず坂下委員。

○坂下委員 どうも唐沢委員、ありがとうございます。消費は基本的に人間の尊厳とか自由を守っている一つの形態だと思っております。そのように見たときに、サービスというのを選択する権利というものが経済格差によって変わってくるというのはあると思うのです。社会として否定されたくない、嫌われたくないというのが蔓延している中では、反対しないツールに没頭していくというのは心理学的にもそうなのだとすることが今日よく分かりました。

1つ質問があるのですけれども、私はデジタルの人間なので、認知コントロールというところで立ちどまって考えるというのがあるではないですか。下のほうにインターフェース設計の例が出ていて、この立ちどまって考えるインターフェース設計は例えばどんなものが議論されているかというのがありますか。

○唐沢委員 私はあまりその点には詳しくないので、知っている範囲のみで申し上げます、すぐに反応を返さないとか、むしろ何か要求を出したときに、AIの側がこちらが出した要求と異なる視点も出すことを促すプロンプトを出すくらいはありそうと思いますし、実際に私もそういう反応を生成AIにされたことがあります。ただ、申し訳ないのですけれども、それぐらいしか私は存じ上げないです。

○坂下委員 分かりました。だから、カードで決済しますというボタンを押したら本当に買っていいかみたいなメッセージが出てくるとか、そういう感じですね。

○唐沢委員 例えばそういうものはあり得るかなと。

○坂下委員 よく分かりました。ありがとうございます。

以上です。

○小塚座長 ありがとうございます。

それでは、馬籠委員、お願いします。

○馬籠委員 お疲れさまでございます。ありがとうございます。非常に勉強になりました。

質問を2点させてください。10ページ目ですが、これは認知コントロールがされていった先といいますか、リアルな人間関係に何かしら悪影響を与えてしまうようなことが生じるのかなと考えております。AIにいつでも質問ができますみたいな状況になったときに、リアルな人間関係とはこれまた別の状態ですので、要はその人たちに頼らなくてもよくなるという状況ができます。これまでは人間関係を頼って生きてきたという部分があると思うのですが、今後は人間関係に頼らなくても生きていくことができそうだとした場合、何かしらリアルな人間関係に与えてしまう影響というのは考えられそうですかというのが1点目です。

もう一点、言い方は難しいのですが、AIに依存していった先に出てくる問題点としては何が考えられそうかなんてことを思っております。例えばAIをつかさどるAIの思想みたいなところに依存しやすくなってしまう。要はAIによる考え方に影響されやすくなるのか、それとも、AIは多分反応率を見てその人の思想を強化していくような側面があると思いますので、そちらのほうに引っ張られていくのか、何か御意見がもしございましたら教えてくださいと幸いです。

○唐沢委員 ありがとうございます。

1点目のほうは、消費者という文脈を外したとしても、AIというものによってリアルな関係がどう影響されるかは、AIに関しての科学技術ELSIの検討の中では議論されていることだと思います。リアルな人間関係が阻害される、それが不要になるといいますか、人間関係に対しての個人の持っているニーズをAIが満たすことによって、リアルな人間関係を維持する必要が減ることは議論されている。ただ、AIがリアルな人間関係をより豊かにするために働きかけるという可能性もあり得ますし、そういう方向で技術活用していく試みもなされているので、リアルな関係が阻害される方向にのみ進むという話にはなっていないように思います。

ただ、そこでのポイントは、リアルな人間関係を私たちが必要としているという前提があると思うのですけれども、それを維持するためにAIをどう使うのかとか、そこにAIをかかわらせることに対してのビジョンがないと難しいかなということ。

それから、人間関係に関しての欲求というのは人々の間でかなり違うので、例えば自己

評価がネガティブであったり、どちらかということと人との関わりが苦手な人に対して、AIを使ってリアルな相互作用を促進する働きかけをすると、かえって不幸になるというか、自分により結果をもたらさないと思われてしまうこともあります。また一方で、そういう人に対して、AIがコンパニオンとして存在するということでポジティブな影響がある、リアルな人間関係を代替するものとしての良い効果もあるという議論もされています。影響はあるのですが、そのバリエーションが大きいのが1点目の答えです。

それから、2点目なのですけれども、何を強化していくかということなのですが、私もAIの仕組みをすごく理解しているわけではないのですが、イメージとしてはより固めていくというか。ある反応をより効果的に引き起こす方向に進むベクトルがあるように思いますので、個々人の好みを強化する、守っていく。それをより効果的に実現していくような可能性がありそうですし、メタな思想がもしあったとしたときに、そのメタな思想というのは他の様々な人からつくられた集約としてのメタな思想だと思えるので、要するに多数派とか平均みたいなことですかね。そういうものに集約していくような可能性もあると思います。

なので、多くの人々がどうなのかを踏まえて、私個人を把握して、その中で効率的に動くことを可能にするという観点からは、よいことなのかもしれませんが、とっぴもないこととか、大きく他と違うことはどうなのか。他とは異なることの持つ可能性を生かすことを組み込むのは、恐らくAIの設計上可能だと思うので、それを意図的に入れて、多数派の偏りや平均化を防ぐことも科学技術ELSIの分野で議論されていたと思います。今、御質問いただいた内容に即して言いますと、問題として起こり得るということは御指摘のとおりですけれども、それを避けることについてもまた一定の議論が今進んでいるのかなというような感触は持っております。

○馬籠委員　ありがとうございます。

よく理解できました。複雑な世界がAIにより平均的になってくるのではないかな、という感想を抱いていて、この例がいいか分からないですけれども、もともとの日本人的みたいなものが、AIによってどんどんスタンダードな何かに変わっていったりするか、なんてことを思っていた次第でございました。

○唐沢委員　ありがとうございます。

○小塚座長　ありがとうございました。

そのほか。

岡崎委員、お願いします。

○岡崎委員　AIを使ってレポートを書いたり、プログラミングをAIに任せたりみたいな人間の心理が整理されたような感じがいたしまして、ありがとうございました。

AIの安全性は有用性と有害性の2軸で測定されることがあるのですけれども、有用性が高いだけでも依存性が高まるといった御示唆だと思っていまして、その場合にAI開発で気をつけるべきことというのはどういったことになるのでしょうかというのが質問です。今回

の話は利用者側の人間としても気をつけることがあるということだと思うのですけれども、AIとして作るときに気をつけるべきことがありましたら教えていただけますと幸いです。

○唐沢委員 ありがとうございます。

それは難しい御質問で、簡単なお答えというのがなかなかできないと思うのですけれども、依存という言葉自体がネガティブなニュアンスを含んでおり、そのことを軸にして言うと、ネガティブな人間の在り方としての依存とはどういう状態かを定義したうえで、開発者の方から見たときに、どういう技術開発、AIの在り方が、その依存状態とつながりうるのかについての知見が恐らく必要だと思います。これは多分開発者の方だけではなくて、AIが人間に与えるインパクトについて検討する研究者との共同で検討することで、その結果を開発者の方々にフィードバックしたときに、望ましくない状態を避けることについて、どこまで技術で対応できるのと考えられるのか。私は人々に与えるインパクトのほうを見ている研究者なので、開発側についてはよく分からないのですけれども、こういう設計や技術なら、こういう人々の反応につながるとか、こういう人々の反応をより促すような設計があると開発側から提案があれば、そのやり取りが有益な議論として成り立つと思うのです。

そういう回りくどい営みでないと、先生の今おっしゃったような問題に対応するような話というのはできないのではないかなと思っているのです。というのは、今、先生がおっしゃった話というのは本当にAI開発のELSIの話そのもののポイントになってしまうので、単純にハウツーでなかなかお答えをお示しすることは難しい御質問かなと思っています。○岡崎委員 ありがとうございます。私もどうしたらいいのかなと思っています、難しい質問を聞いてしまいました。

○唐沢委員 でも、非常に大事な点で、むしろ開発者側の先生方がそういうことを大事だと思ってくださって、それをどういうふうにしたらいいかという議論のプラットフォームに乗っていただいて議論を進めていくということが大事だと思いました。ありがとうございます。

○岡崎委員 ありがとうございます。

○小塚座長 ありがとうございました。

岡崎委員はむしろ開発者側に近いお立場だと思いますけれども、まさに御指摘になったようなAIの有用性自体がある種の人間の行動、あるいは個別の人間はそれで満足しているとしても、社会全体に弊害があるときに、AIの開発プロセスの中でそこにガードレール的なものを入れていくというのは技術的には可能性はあると理解してよろしいですか。

○岡崎委員 例えばこの質問をされたけれども、これは答えると人間にとってよくないなとか成長にならないなみたいな判断というのはあると思うのです。でも、それは一方で有用性を下げるということでもあるので、そこは難しいよなとは思っています。

○小塚座長 ありがとうございます。逆に言うと、有用性をコントロールすることでより大きな弊害に対処するということですね。

もう一つの方向性は先ほど坂下委員の御質問で出てきましたけれども、AIそのものではなくて、それを使うインターフェースのほうでコントロールしていく。この方向性はもう一つあるということだと思います。

ありがとうございます。まだまだお聞きしたいこともいろいろあるのですけれども、そろそろ時間になってまいりましたので、本日は一旦質疑応答と意見交換はこの辺りとさせていただきますらと思います。またどんどん考える材料が増えてまいりますので、これを踏まえて私ども議論をしていきたいと思っています。

唐沢委員、本当に御主張先でお忙しい中からお話をいただきまして、どうもありがとうございました。

○唐沢委員 どうもありがとうございました。

○小塚座長 それでは、事務連絡を事務局のほうからいただけますでしょうか。

《3. 閉会》

○江口企画官 次回の本専門調査会につきましては、確定次第、事務局より御連絡させていただきます。

以上です。

○小塚座長 ありがとうございました。

それでは、本日はこれにて閉会です。お忙しい中、お集まりいただきましてありがとうございました。

以 上